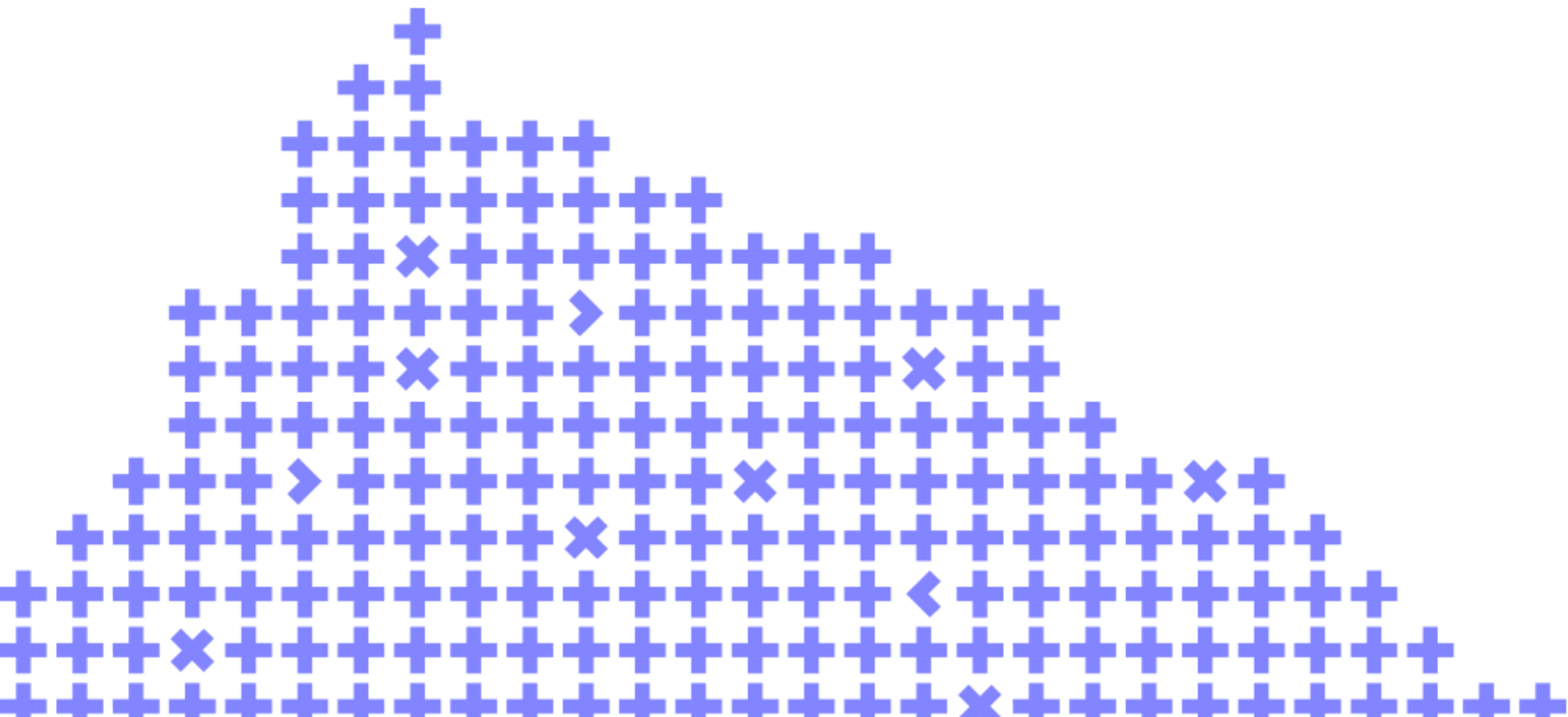


You need Cloud to manage Cloud: Kubernetes as best way to manage OpenStack cloud

Vadim Ponomarev



Co-organizer

Yandex

What Is OpenStack?

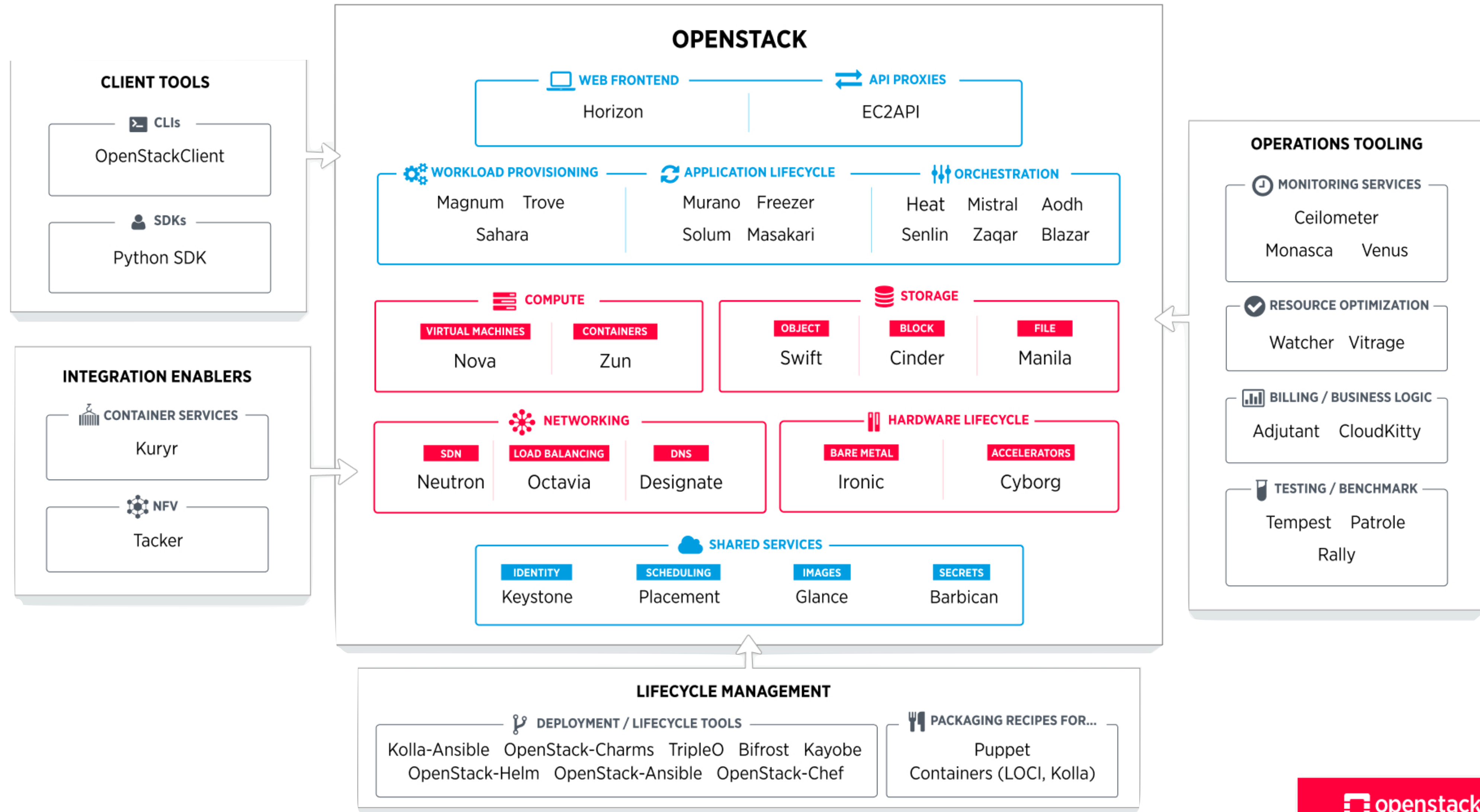
- **open-source cloud computing platform**
- **created by Rackspace and NASA in 2010**
- **written in python**
- **modular and microservices architecture**
- **used for a public/private cloud**



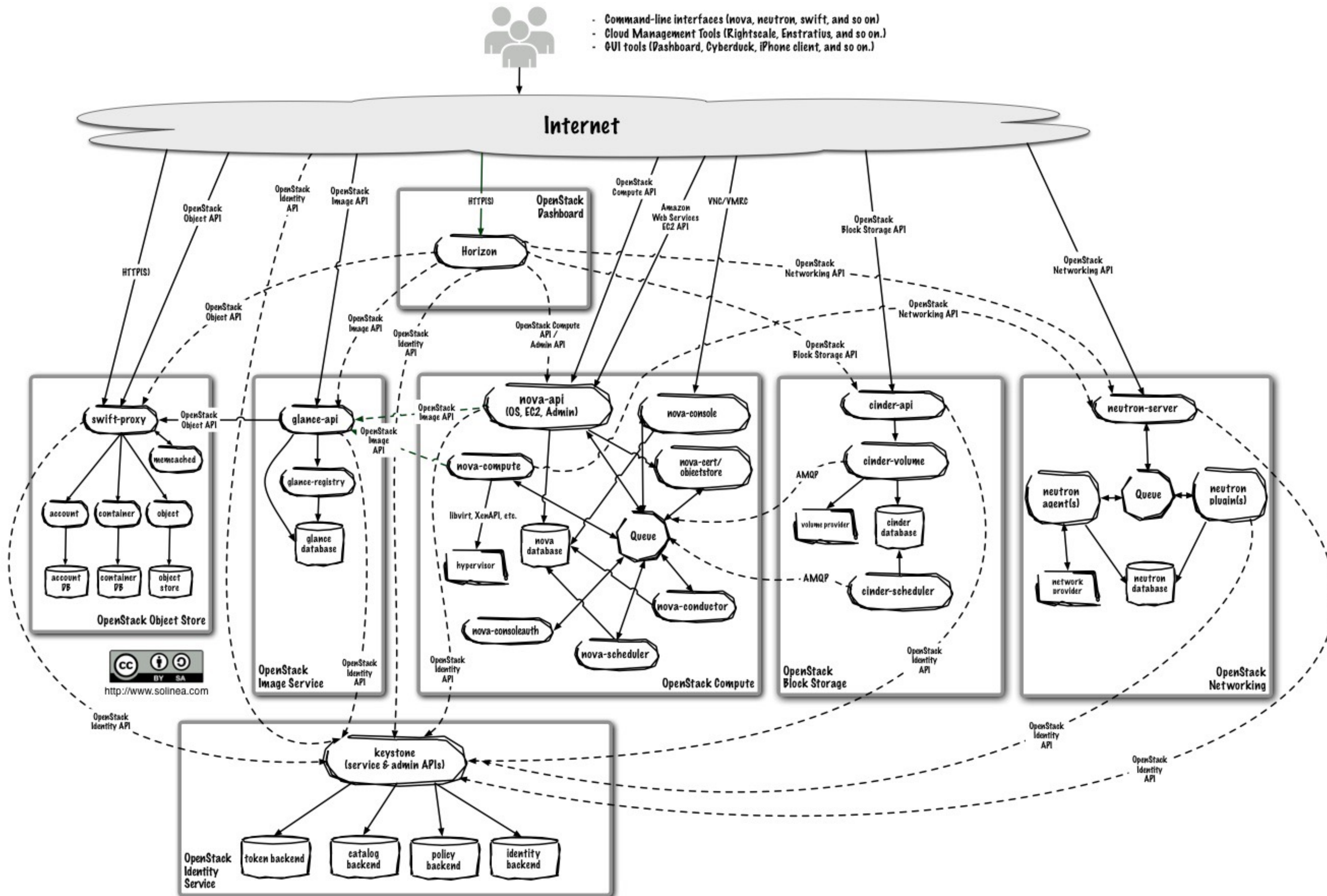
Why do we need OpenStack?

- **open-source self-hosted solution for private / public clouds**
- **Vmware alternative with zero price tag**
- **a strong network isolation**

OpenStack architecture



OpenStack architecture



What's the problem?

- **hundreds of microservices**
- **hundreds of bare-metal servers**
- **a huge python codebase**
- **a full update at least twice a year (upstream release period)**

WHAT COULD

POSSIBLY GO WRONG?



Why Kubernetes?

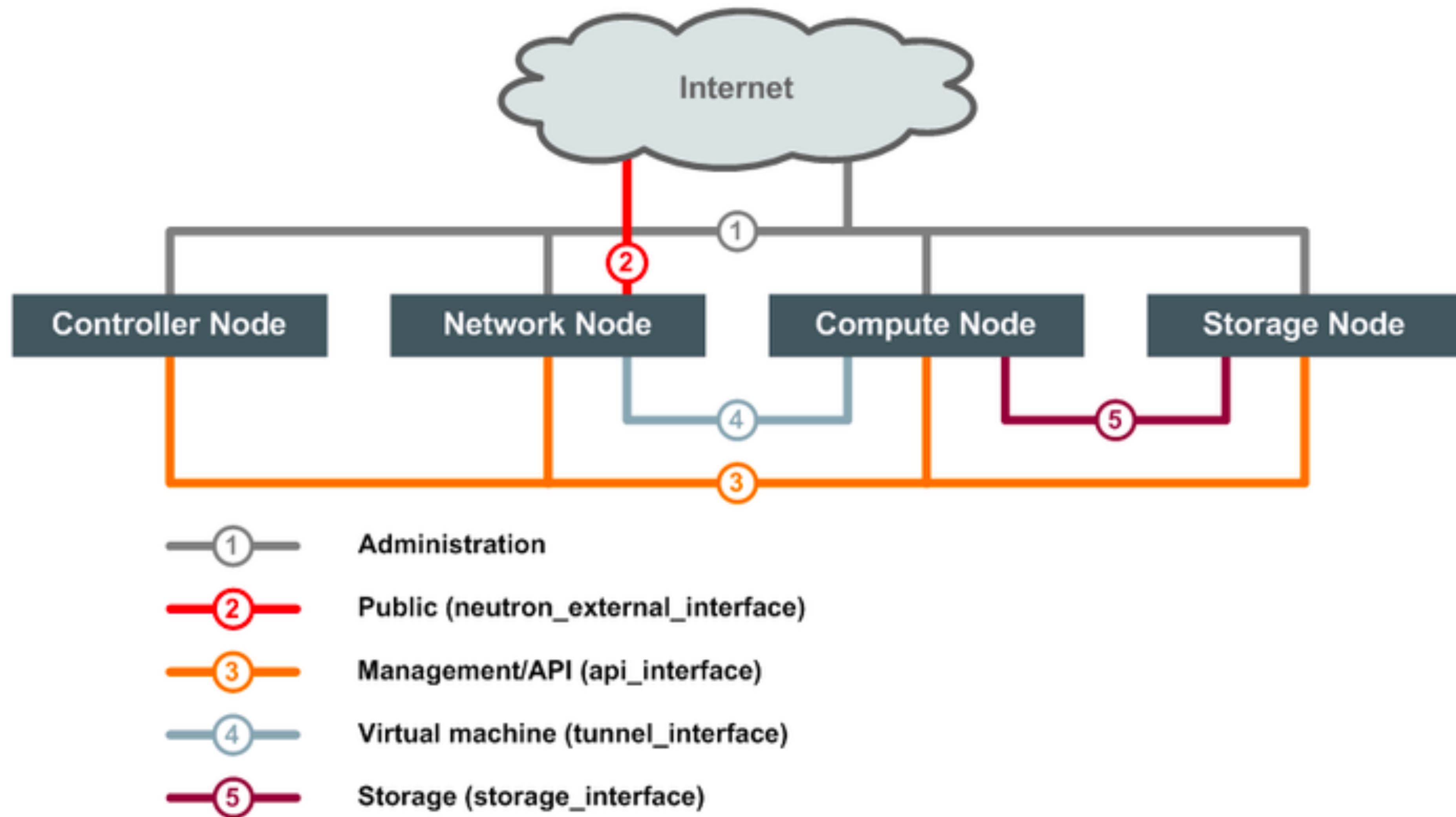
- ✓ built to manage thousands of microservices
- ✓ can scale to hundreds of nodes
- ✓ containerization solves the problem with dependencies
- ✓ self-healing, high availability, healthchecks
- ✓ and many other benefits ...

But ...





- **OpenStack is not just an application**
- **VMs will be running on k8s workers**
- **OpenStack has its own network stack**
- **a complicated order of starting services**
- **a storage based on Ceph**

Moreover



And how does K8s help here?

General Tips

- do not reinvent the wheel
- use openstack-helm 
- use official docker images when possible 
- run all the OpenStack services in one namespace
- RTFM (if you can find it)



Database

- **Percona XtraDB Cluster with k8s operator**
- **separate database cluster for Neutron (network system)**
- **use fast SSDs if cluster > 50 compute nodes**
- **monitoring**

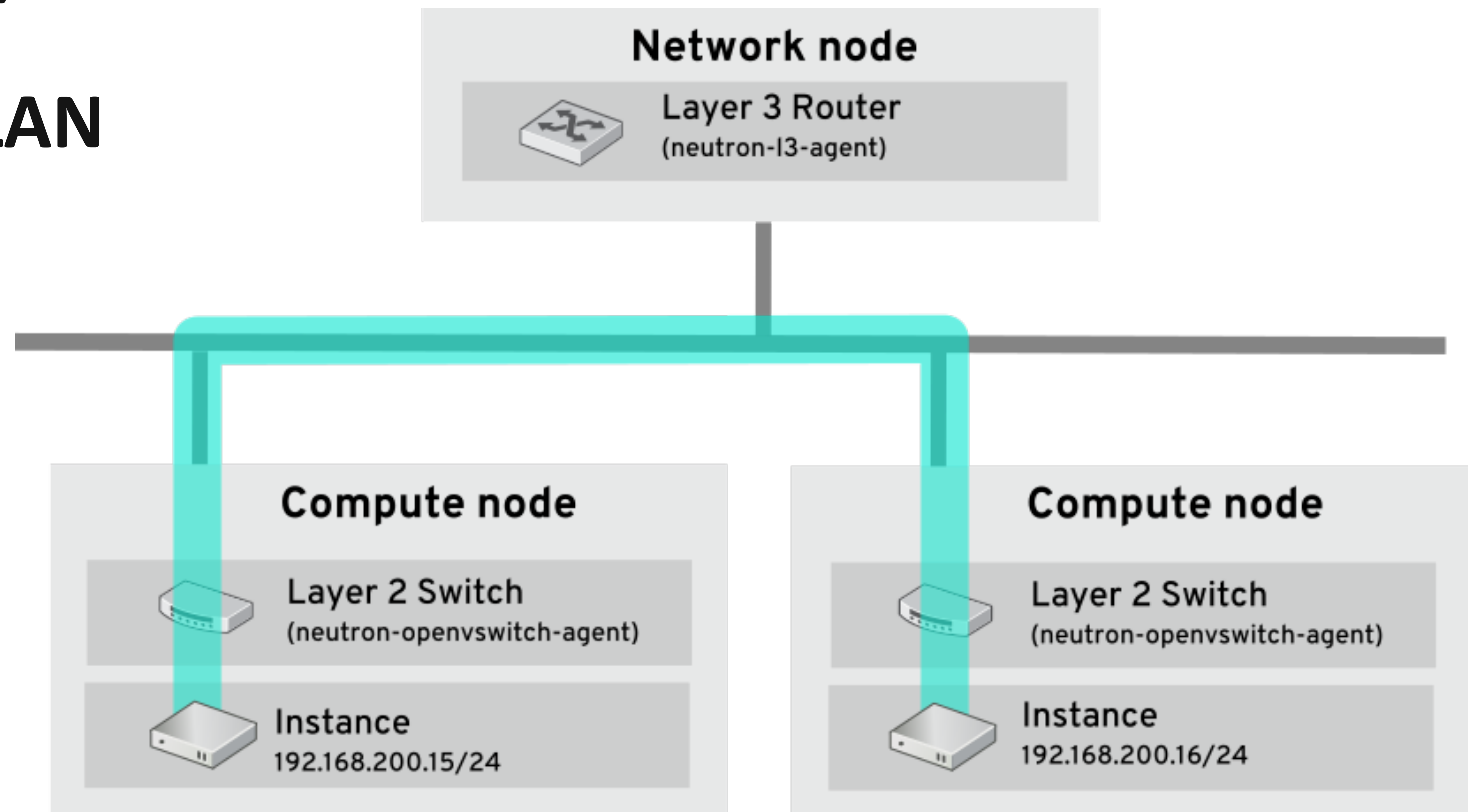
Storage system

- **CephFS is the most popular**
- **one Ceph cluster for k8s and for OpenStack (different pools)**
- **a separate physical network for a storage**
- **dedicated storage hosts if you have the budget:**
 - + to reduce load**
 - + to reduce chances of losing data**
 - + to have faster reboot of compute nodes**

How OS network works

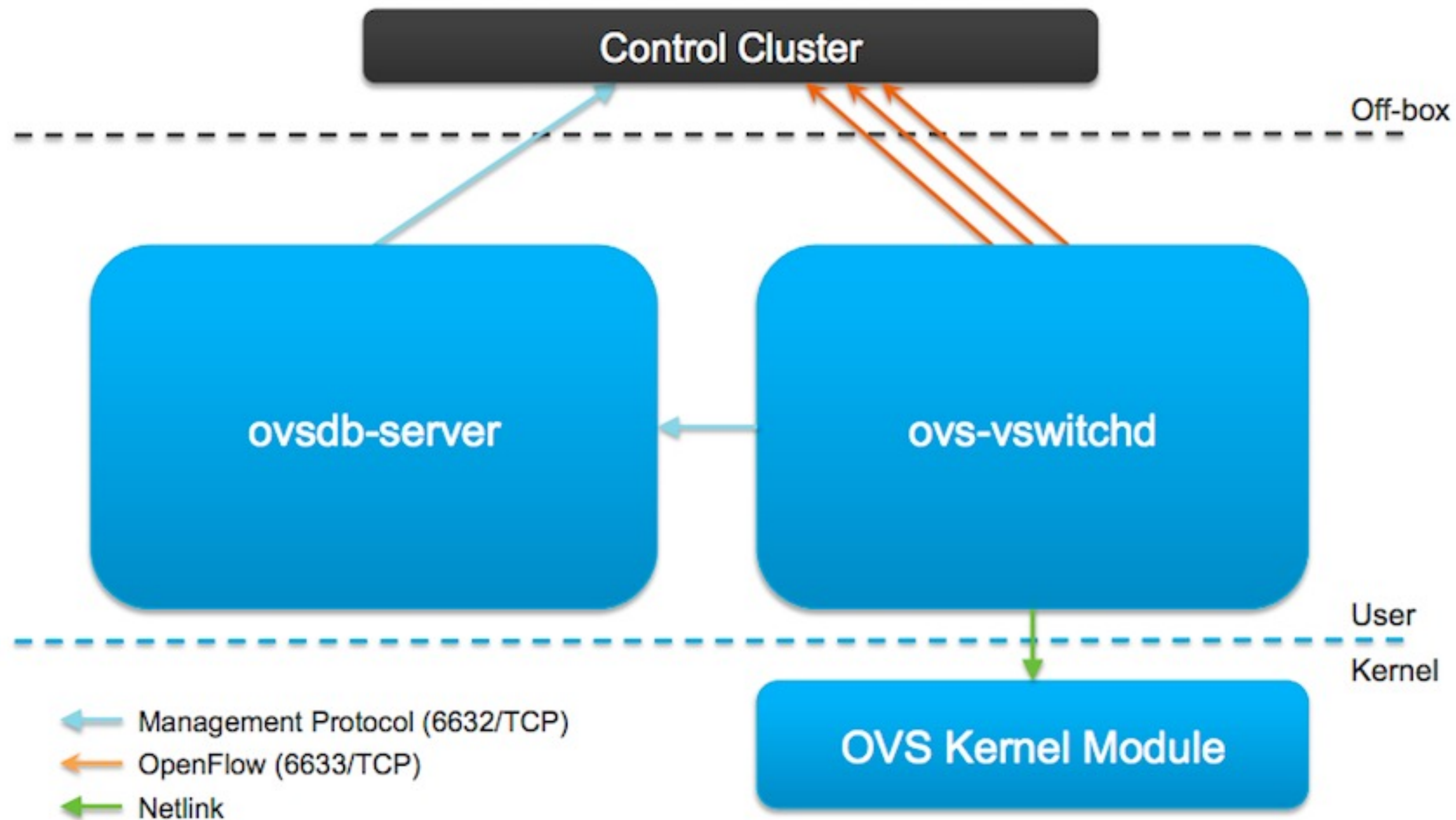


- SDN OpenvSwitch / OVN
- L2: VXLAN / Geneve / VLAN
- L3: virtual routers / OVN
- dnsmasq DHCP / DNS
- service called “Neutron”



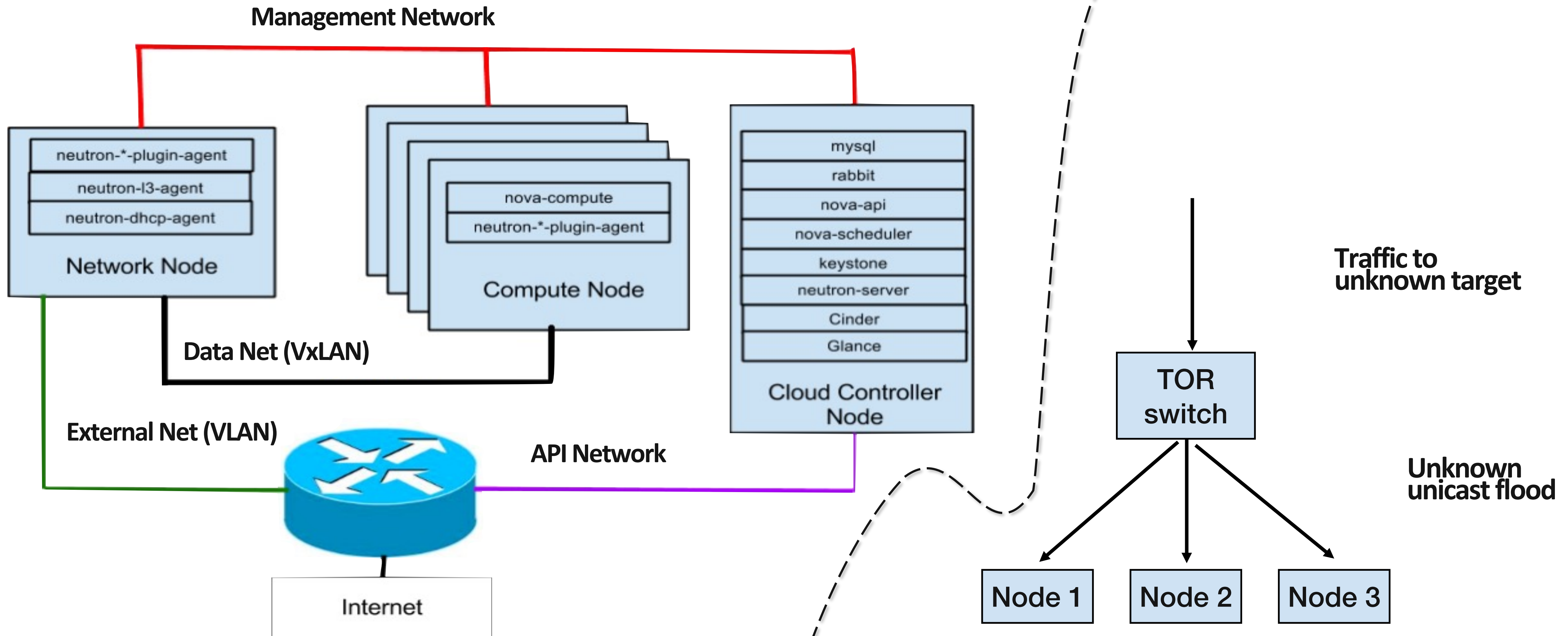
Network challenges

1. How does OpenvSwitch/OVN configure host system?



Network challenges

2. External networks only VLAN based



Network Tips: OVS

1. OpenvSwitch daemon:

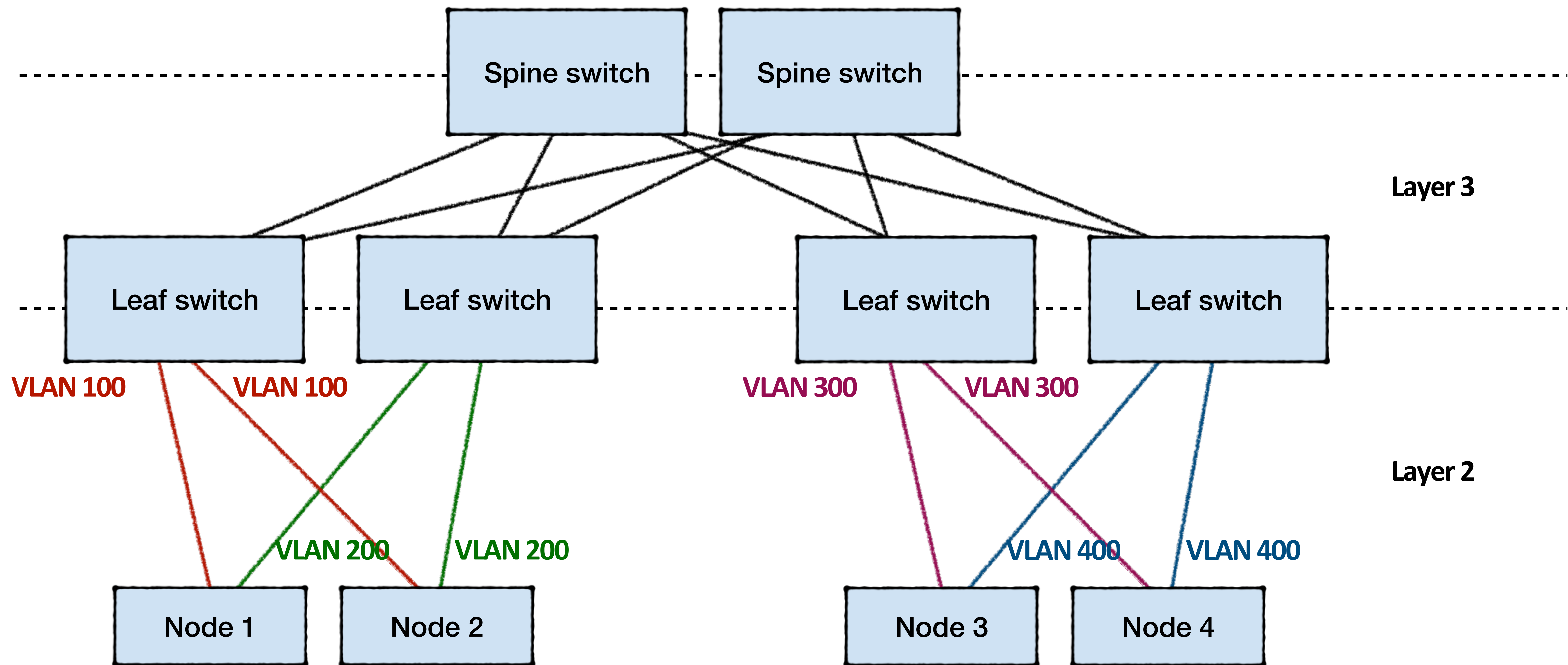
- host network
- capabilities
- run as root
- mount */run* directory from the host system



```
spec:
  hostNetwork: true
  ...
  containers:
    ...
    securityContext:
      capabilities:
        add:
          - NET_ADMIN
          - SYS_MODULE
          - SYS_NICE
      readOnlyRootFilesystem: true
      runAsUser: 0
    ...
    volumeMounts:
      - mountPath: /run
        name: run
  ...
  volumes:
    - hostPath:
        path: /run
        type: Directory
      name: run
```

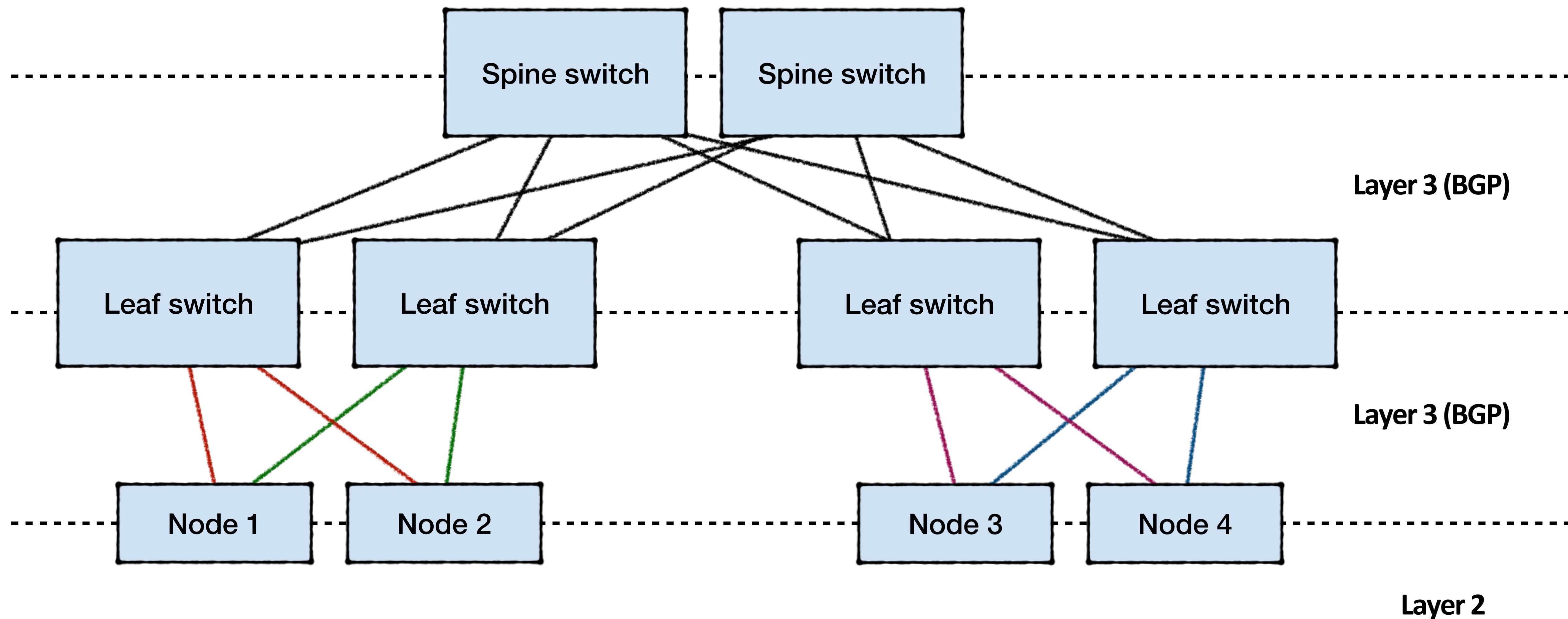
Network Tips: S&L TOR VLANs

2. External networks:



Network Tips: S&L without VLANs

2. External networks:



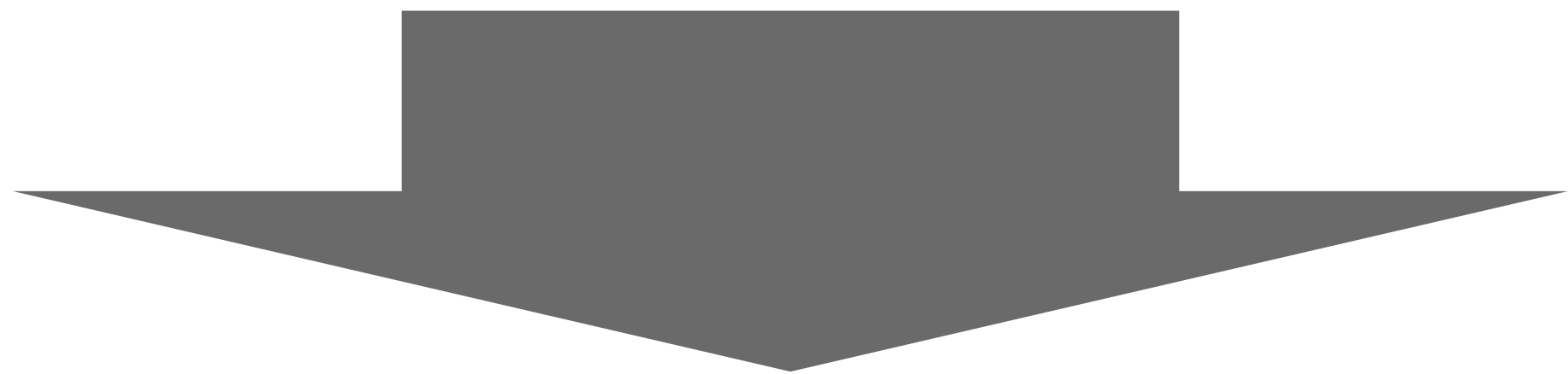
Network Tips: solutions

- use segments extension and per-rack VLANs
- use BGP dynamic routing plugin
- use DVR routers when it's possible
- use EVPN-VXLAN network in the data center



Compute

- **Nova configures KVM on the host system**
- **VM can have a direct access for network/GPU cards**

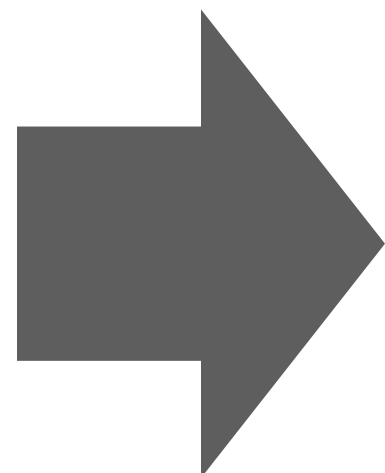


- **Privileged libvirt container**
- **Mounts from the host system:**
 - **/lib/modules**
 - **/var/lib/nova**
 - **/var/lib/libvirt**
 - **/run**
 - **/sys/fs/cgroups**

Compute

- State directories with RW access from all the hosts for the migrations

```
spec:
...
  securityContext:
    privileged: true
...
  - mountPath: /var/lib/nova-cephfs
    name: cephfs-var-lib-nova
...
  volumes:
  - name: cephfs-var-lib-nova
    persistentVolumeClaim:
      claimName: var-lib-nova
```



```
# k describe pvc -n openstack var-lib-nova
Name:          var-lib-nova
Namespace:     openstack
StorageClass:  cephfs
...
Access Modes:  RWX
VolumeMode:    Filesystem
Mounted By:    libvirt-libvirt-default-5v2ld
               libvirt-libvirt-default-7ckld
               libvirt-libvirt-default-7prvw
               libvirt-libvirt-default-f2vpp
               libvirt-libvirt-default-sq5rs
               nova-compute-default-2kmhd
               nova-compute-default-4cpf8
               nova-compute-default-728lq
               nova-compute-default-r5lk7
               nova-compute-default-rtxkm
```

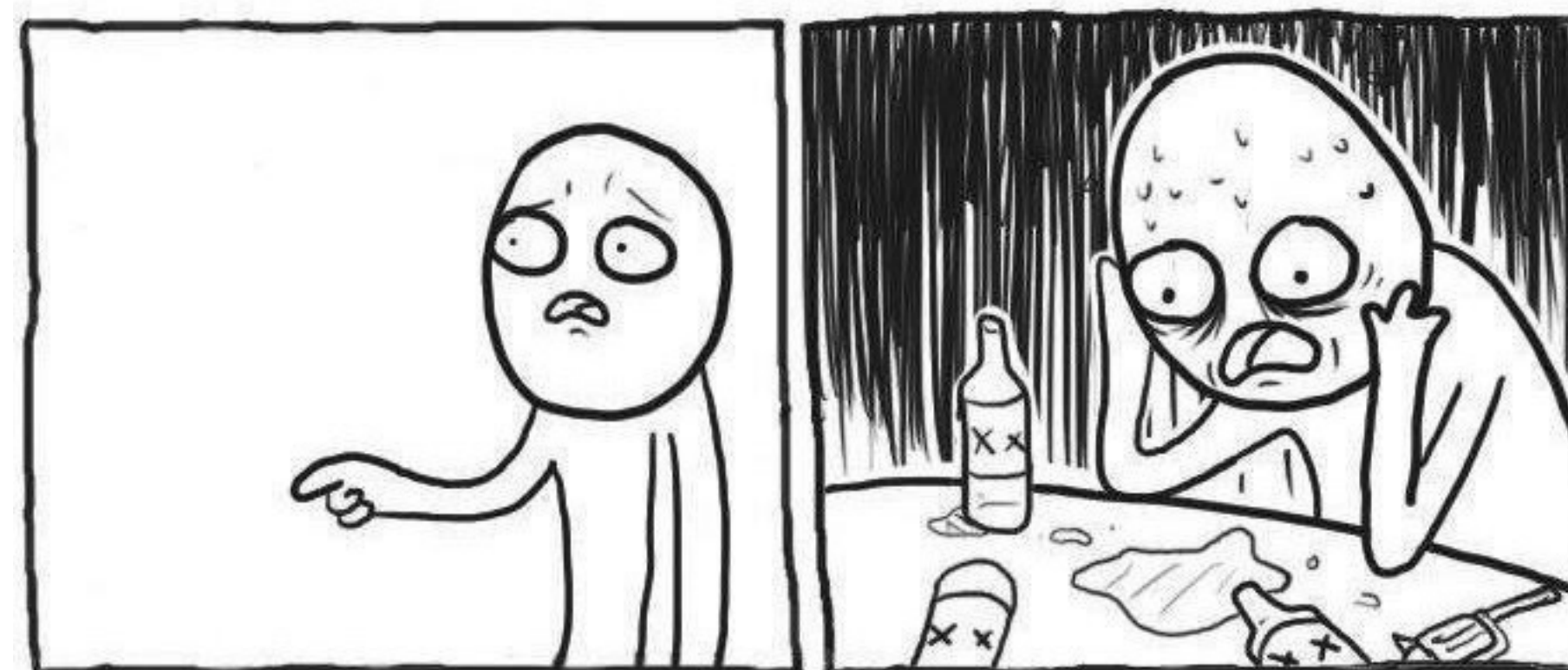
Is OpenStack ready?

- **Bad or non-existent healthchecks**
- **No graceful restart**
- **Multiline logs (no json support!)**

```
2022-11-18 00:00:40.407 52 ERROR neutron.agent.dhcp.agent Traceback (most recent call last):
2022-11-18 00:00:40.407 52 ERROR neutron.agent.dhcp.agent   File "/var/lib/openstack/lib/python3.6/site-packages/neutron/agent/dhcp/agent.py", line 1044, in _report_state
2022-11-18 00:00:40.407 52 ERROR neutron.agent.dhcp.agent     ctx, self.agent_state, True)
2022-11-18 00:00:40.407 52 ERROR neutron.agent.dhcp.agent   File "/var/lib/openstack/lib/python3.6/site-packages/neutron/agent/rpc.py", line 101, in report_state
2022-11-18 00:00:40.407 52 ERROR neutron.agent.dhcp.agent     return method(context, 'report_state', **kwargs)
2022-11-18 00:00:40.407 52 ERROR neutron.agent.dhcp.agent   File "/var/lib/openstack/lib/python3.6/site-packages/oslo_messaging/rpc/client.py", line 181, in call
2022-11-18 00:00:40.407 52 ERROR neutron.agent.dhcp.agent     transport_options=self.transport_options)
2022-11-18 00:00:40.407 52 ERROR neutron.agent.dhcp.agent   File "/var/lib/openstack/lib/python3.6/site-packages/oslo_messaging/transport.py", line 129, in _send
2022-11-18 00:00:40.407 52 ERROR neutron.agent.dhcp.agent     transport_options=transport_options)
2022-11-18 00:00:40.407 52 ERROR neutron.agent.dhcp.agent   File "/var/lib/openstack/lib/python3.6/site-packages/oslo_messaging/_drivers/amqpdriver.py", line 682, in send
2022-11-18 00:00:40.407 52 ERROR neutron.agent.dhcp.agent     transport_options=transport_options)
2022-11-18 00:00:40.407 52 ERROR neutron.agent.dhcp.agent   File "/var/lib/openstack/lib/python3.6/site-packages/oslo_messaging/_drivers/amqpdriver.py", line 672, in _send
2022-11-18 00:00:40.407 52 ERROR neutron.agent.dhcp.agent     raise result
2022-11-18 00:00:40.407 52 ERROR neutron.agent.dhcp.agent oslo_messaging.rpc.client.RemoteError: Remote error: InternalError (pymysql.err.InternalError) (1047, 'WSREP has not ye
2022-11-18 00:00:40.407 52 ERROR neutron.agent.dhcp.agent (Background on this error at: http://sqlalche.me/e/2j85)
```


Is OpenStack ready?

- **Bad monitoring abilities**
- **Complex dependencies between components**
- **Difficult to customize images with components**



If everything is so bad, why K8s?

- **Anyway, it gives better control over hundreds of services with K8s**
- **It gives more stability with updates**
- **Self-healing, HA, isolation, etc.**
- **It's easier to control at a large scale**
- **K8s is more popular than OpenStack**

Leave your feedback!

You can rate the talk and give feedback on what you've liked or what could be improved

 <https://www.linkedin.com/in/v-pon/>

 @velizarx

 <https://github.com/velp>



Co-organizer

Yandex